



Published in final edited form as:

Biochem Biophys Res Commun. 2014 April 18; 446(4): 1067–1072. doi:10.1016/j.bbrc.2014.03.060.

Excision of 5-hydroxymethylcytosine by DEMETER family DNA glycosylases

Hosung Jang^a, Hosub Shin^a, Brandt F. Eichman^b, and Jin Hoe Huh^{a,*}

^aDepartment of Plant Science, Research Institute for Agriculture and Life Sciences, and Plant genomics and Breeding Institute, Seoul National University, Seoul 151-921, Korea

^bDepartment of Biological Sciences, Vanderbilt University, Nashville, TN 37235, USA

Abstract

In plants and animals, 5-methylcytosine (5mC) serves as an epigenetic mark to repress gene expression, playing critical roles for cellular differentiation and transposon silencing. Mammals also have 5-hydroxymethylcytosine (5hmC), resulting from hydroxylation of 5mC by TET family-enzymes. 5hmC is abundant in mouse Purkinje neurons and embryonic stem cells, and regarded as an important intermediate for active DNA demethylation in mammals. However, the presence of 5hmC in plants has not been clearly demonstrated. In *Arabidopsis*, the DEMETER (DME) family DNA glycosylases efficiently remove 5mC, which results in DNA demethylation and transcriptional activation of target genes. Here we show that DME and ROS1 have a significant 5hmC excision activity *in vitro*, although we detected no 5hmC in *Arabidopsis*, suggesting that it is very unlikely for plants to utilize 5hmC as a DNA demethylation intermediate. Our results indicate that both plants and animals have 5mC in common but DNA demethylation systems have independently evolved with distinct mechanisms.

Keywords

DNA demethylation; DEMETER; DNA glycosylase; 5-hydroxymethylcytosine; 5-methylcytosine

1. Introduction

DNA methylation is a simple but important epigenetic modification, playing an important role for diverse biological processes such as transcriptional regulation, cellular differentiation, gene imprinting, and silencing of transposable elements [1-4]. In higher eukaryotes, DNA methylation often refers to DNA methyltransferase-catalyzed methylation of the C5 position of cytosine, to generate 5-methylcytosine (5mC).

DNA methylation can be reversed by either passive or active mechanisms [5]. Passive DNA demethylation is replication-dependent and occurs gradually when maintenance DNA methylation is suppressed in dividing cells. By contrast, active DNA demethylation takes place in a replication-independent manner, and requires certain enzyme activities. Recent

*Corresponding author Address: Department of Plant Science, Seoul National University, Seoul 151-921, Korea; Phone: +82-2-880-4562; huhjh@snu.ac.kr.

studies suggest that two different mechanisms may operate for active DNA demethylation. One involves a direct removal of 5mC from DNA, which is consequently replaced with unmethylated cytidine via the base excision repair (BER) pathway. The other implicates an enzymatic modification of 5mC to other bases, which do not require direct excision but nullify the silencing effect of 5mC. Recently, the latter has emerged as an important DNA demethylation pathway in mammals, in which the ten-eleven translocation (TET) family of proteins catalyze the oxidation of 5mC to 5-hydroxymethylcytosine (5hmC) [6-12]. Even though the conversion of 5mC to 5hmC may cancel the effect of DNA methylation in part, further TET-dependent oxidation of 5hmC produces 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC), both of which can be excised by thymine-DNA glycosylase (TDG) initiating the BER pathway for complete DNA demethylation [7,13].

It is believed that plants and animals have evolved distinct DNA demethylation systems, particularly in that plants have enzymes that are able to remove 5mC from DNA [4,14]. The DEMETER (DME) family of proteins are DNA glycosylases that primarily recognize and excise 5mC from DNA [15-19]. DME was initially identified in *Arabidopsis* as a transcription activator of *MEDEA* PcG gene [20], whose maternal-specific expression in fertilized endosperm was accompanied with reduced DNA methylation at the promoter region [16]. In *Arabidopsis*, the DME family includes REPRESSOR OF SILENCING 1 (ROS1), DME-LIKE 2 (DML2), and DML3 [20,21]. As bifunctional DNA glycosylases, the DME family proteins catalyze the cleavage of an N-glycosidic bond between the base and the ribose sugar, and their additional AP lyase activity induces a DNA strand break via the β -elimination process [15-19]. It is likely that BER enzymes subsequently participate in the DNA demethylation process by incorporating unmethylated cytidine in place of excised 5mC [3,4,14]. Consequently, the DME family-initiated BER pathway is regarded as a predominant DNA demethylation mechanism in plants. However, an alternative active DNA demethylation pathway involving 5mC modification as in mammals has not yet been reported in plants.

In this study, we investigated base excision activity of DME and ROS1 for 5mC oxidation derivatives such as 5hmC, 5fC, and 5caC. Despite being less preferred than 5mC, the 5hmC base was found to be significantly excised by DME and ROS1 *in vitro*, suggesting the possibility of alternative route of DNA demethylation in plants. Additionally, we explored the presence of 5hmC in the *Arabidopsis* genome in order to understand whether 5hmC excision activity of DME/ROS1 is biologically relevant and associated with DNA demethylation dynamics [22], or simply reflects a broad substrate specificity range of this family of proteins [23].

2. Materials and methods

2.1. Cloning of DME and ROS1

The DME N677 IDR1::lnk fragment (hereafter called “DME ”, [21]) in which both N-terminal 677 amino acids and IDR1 were removed from DME, was cloned into the *Bam* HI and *Sal*I sites of the pLM302 vector (Center for Structural Biology, Vanderbilt University). The ROS1 N509 fragment (hereafter called “ROS1 ”) was PCR-amplified from full-length ROS1 cDNA and then introduced into the pLM302 vector at the sites of *Eco* RI and *Sal*I.

Both DME and ROS1 were fused with 6xHis and maltose binding protein (MBP) tags at the N-terminus.

2.2. Protein expression and purification

DME and ROS1 fused with 6xHis + MBP tags were expressed in *E. coli* Rosetta2 (DE3) strains (Novagen). Protein purification steps were essentially the same as described by Mok et al. [21]. Briefly, 6xHis-MBP-DME or -ROS1 was sequentially purified through a HisTrap FF column (GE Healthcare) and a HiTrap Heparin HP column (GE Healthcare), and gel filtration was performed on a HiLoad 16/60 Superdex 200-pg column (GE Healthcare) (Suppl. Fig. 1).

2.3. In vitro base excision assay for 5mC derivatives

Thirty five-mer oligonucleotides containing cytosine, 5mC, 5hmC, 5fC, and 5caC were purchased from Midland Certified (TX, USA) (Suppl. Table 1). Forty pmol of each oligonucleotide was radiolabeled with [γ -³²P]ATP (6,000 Ci/mmol, Perkin Elmer Life Sciences) using T4 polynucleotide kinase (Takara) and then annealed to a complementary oligonucleotide to produce double-stranded DNA substrate. *In vitro* base excision assay was performed as previously described [21]. Briefly, 25 nM of each radiolabeled oligonucleotide substrate was incubated with 100 nM of MBP-DME or 85 nM of MBP-ROS1 in the glycosylase reaction buffer (10 mM Tris-HCl, pH 7.4, 50 mM NaCl, 0.5 mM dithiothreitol, 200 μ g/mL BSA) at 37 °C for 1 hr. Reactions were terminated by adding an equal volume of stop solution (98% formamide, 10 mM EDTA, 0.2% xylen cyanol FF, 0.2% bromophenol blue) and heat-denaturing at 95 °C for 10 min. Reaction products were separated on a 15% denaturing polyacrylamide gel containing 7.5 M urea and 1 \times TBE.

2.4. Kinetics analysis

Twenty five nM of oligonucleotide substrate containing 5mC or 5hmC was incubated with excess amount of DME in a time course manner (0.5, 1, 2, 3, 5, 10, 25, and 60 min). Reaction was terminated by adding 100 mM NaOH and boiling for 10 min. Reactions were denatured and separated on a 15% denaturing polyacrylamide gel. The gel was exposed to a phosphorimager screen (Fujifilm) and the radioactivity was measured using the Fujifilm BAS-5000 phosphorimager.

2.5. Substrate competition assay

To produce the unlabeled competitor substrates with the same concentration of radiolabeled substrates, 5mC- and 5hmC-containing oligonucleotides were purified with the QIAquick Nucleotide Removal Kit (Qiagen) and annealed with complementary oligonucleotides. One hundred nM of MBP-DME was incubated with 25 nM of radiolabeled oligonucleotides in the presence of 5mC or 5hmC containing competitor substrates. The amount of competitor oligonucleotides varied from 0 to 20-fold (0, 125, 250, 500 nM) over radiolabeled substrates. After adding an equal volume of stop solution and boiling for 10 min, reactions were separated on a 15% denaturing polyacrylamide gel and analyzed with phosphorimager scanning as described above.

2.6. In vitro nucleotide incorporation assay

Twenty five nM of radiolabeled oligonucleotide substrate containing 5mC or 5hmC was reacted with 100 nM of MBP-DME and 10 U of Endonuclease IV (NEB) in the glycosylase reaction buffer at 37 °C for 1 h. Following heat-inactivation at 65 °C for 15 min, the reaction was subjected to nucleotide incorporation with 0.1 mM dCTP using 5 U of Klenow fragment (3' to 5' exo-) (NEB) at 25 °C for 15 min.

2.7. TLC analysis

Genomic DNA was isolated from young leaves, and floral buds of *Arabidopsis thaliana* Columbia ecotype using the standard CTAB method. Control DNA was prepared by PCR amplification using 5-methyl-dCTP or 5-hydroxymethyl-dCTP (Zymo Research) instead of dCTP in the PCR reaction. TLC analysis was performed using slightly modified protocols from Ito et al. [9]. Briefly, 2 µg of genomic DNA or control DNA was digested with 50 U of Taq^αI (NEB) and 5 U of Calf Intestinal Alkaline Phosphatase (NEB). After purification with QIAquick Nucleotide Removal Kit (Qiagen), dephosphorylated DNA was end-labeled with 20 µCi of [γ -³²P]ATP (6,000 Ci/mmol, Perkin Elmer Life Sciences) using T4 polynucleotide kinase (Takara). After purification, radiolabeled DNA was digested with 50 U of Benzonase (Sigma) and 4 mU of Phosphodiesterase I (Sigma) at 37 °C for 1 h. Digested DNA fragments were concentrated and spotted on a PEI-cellulose TLC plate (Merck) and separated for 16 h in the 1D-TLC buffer (isobutyric acid : NH₄OH : H₂O = 33 : 1 : 10). In the second dimension, TLC was performed in the 2D-TLC buffer (isopropanol : HCl : H₂O = 70 : 15 : 15) for 16 h. The plate was exposed to an X-ray film or analyzed by phosphorimager scanning.

3. Results

3.1. DME and ROS1 excise both 5mC and 5hmC in vitro

The DME/ROS1 family proteins are known to primarily excise 5mC [15-19] initiating the BER pathway for DNA demethylation. However, like many other DNA glycosylases, they also recognize and excise a broad range of DNA bases including thymine from a T/G mismatch and several modified bases paired with guanine such as 5-fluoruracil, 5-bromocytosine, 5-bromouracil, and 5-hydroxyuracil [16-18, 24]. This suggests that DME/ROS1 may have excision activity for some 5mC derivatives that are chemically or enzymatically modified. In particular, a few oxidized bases (5hmC, 5fC, and 5caC; Fig. 1A) catalyzed by the mammalian TET1-family proteins immediately drew our attention because they may serve as intermediates leading to TDG-mediated base excision and DNA demethylation even in plants.

Therefore, we prepared radiolabeled-oligonucleotides containing cytosine oxidation derivatives such as 5hmC, 5fC, and 5caC along with 5mC and performed an *in vitro* base excision assay with purified MBP-DME and MBP-ROS1 (Fig. 1B). Because of increased protein stability, we used truncated forms of DME and ROS1 fragments that still retain all conserved domains necessary for *in vitro* base excision activity [21]. We observed that MBP-DME and MBP-ROS1 excised both 5mC and 5hmC, generating 3'-phosphor- α , β -unsaturated aldehyde (3'-PUA) as a major product via a β -elimination process, even though

the latter was less preferentially processed (Fig. 1C). In addition, MBP-DME and MBP-ROS1 were able to excise 5caC (Fig. 1C), despite much less excision activity observed. However, we did not observe any discernable activity of MBP-DME and MBP-ROS1 for 5fC (Fig. 1C). These results suggest that both DME and ROS1 prefer 5mC to other 5mC derivatives and that they have a broad substrate specificity range even though the substituents at C5' of the cytosine ring differ. Notably, the finding that both DME and ROS1 have glycosylase activity for 5hmC suggests not only the possibility of the presence of 5hmC in the plant genome but also an alternative route of active DNA demethylation pathway to remove 5mC.

3.2. Reconstitution of DNA demethylation pathway following 5mC or 5hmC excision

We expect that following 5mC excision by the DME/ROS1 family of proteins, subsequent BER enzymes participate in the DNA demethylation process by incorporating unmethylated cytidine in place of excised 5mC. We reconstituted the *in vitro* BER-mediated DNA demethylation pathway by showing the replacement of 5mC or 5hmC with unmethylated cytidine after DME base excision. When the radiolabeled oligonucleotide containing 5mC was reacted with MBP-DME for 1 h, 3'-PUA and 3'-phosphate were produced via β - and δ -elimination processes, respectively (lane 2 in the left panel of Fig. 1D). A subsequent reaction with AP endonuclease converted both products to 3'-OH (lane 3 in the left panel of Fig. 1D). Finally, the dCTP incorporation at the site of base excision was achieved by Klenow DNA polymerase (lane 4 in the left panel of Fig. 1D). Similar results were obtained when 5hmC-containing oligonucleotide was subjected to the same series of reactions (right panel of Fig. 1D). These results demonstrate that DME has the capacity to initiate active DNA demethylation after 5mC or 5hmC excision via the BER pathway.

3.3. DME has a lower 5hmC excision rate over 5mC

To compare base excision efficiencies of DME for 5mC and 5hmC, a time-dependent kinetics analysis was performed (Fig. 2). In order to measure the formation of single base excision products, reactions were terminated at various time points and treated with strong base (100 mM NaOH) to obtain homogeneous δ -elimination products. Because many glycosylases are known to bind tightly to the reaction products abasic sites, we obtained k_{cat} values under single-turnover conditions with enzyme in excess of substrate (see Materials and Methods for experimental details). The single turnover rate constants (k_{cat-sd}) for 5mC and 5hmC were 0.049 min^{-1} and 0.019 min^{-1} , respectively, indicating that DME has a 2.6-fold higher rate of base excision for 5mC over 5hmC (Fig. 2).

3.4. 5mC is more preferred substrate for DME than 5hmC

In order to assess differential DME substrate preference between 5mC and 5hmC, we performed a substrate competition assay (Fig. 3A). When the radiolabeled oligonucleotide substrates containing 5mC or 5hmC were reacted with MBP-DME in the presence of varying amounts of unlabeled oligonucleotide competitors, the product formation decreased as the amount of cold competitor increased (Fig. 3A). Notably, the product formation for 5mC was more affected by the competitors containing 5mC than 5hmC (left panels of Fig. 3A and 3B). Similar results were obtained when radiolabeled 5hmC oligonucleotides were

competed with 5mC- or 5hmC-containing cold oligonucleotides (right panels of Fig. 3A and 3B). These results indicate that 5mC is a more preferred substrate for DME than 5hmC.

3.5. No 5hmC is detected in plant DNA

The 5hmC excision activity of DME/ROS1 raises the possibility of the presence of 5hmC in the plant genome that can be utilized as another DNA demethylation substrate. In order to explore the existence of 5hmC in plant DNA, we extracted genomic DNA from several tissues in *Arabidopsis* including floral buds where DME is primarily expressed [20]. Isolated genomic DNA was treated with Taq^αI and calf intestine phosphatase and radiolabeled with polynucleotide kinase (Suppl. Fig. 2). After digestion with Benzonase and Phosphodiesterase I, the nucleotides were separated on a TLC plate. In the 1D TLC analysis, compared to the control experiment (Fig. 4A), no spots corresponding to 5hmC were detected in any tissues examined (Fig. 4B). Even in the 2D TLC analysis, a small amount of 5hmC (0.2% of 5hmC approximately equal to its content in mammalian genome [9]) was identifiable (Fig. 4C), but no 5hmC signal was detected in plant tissues (Fig. 4D and Suppl. Fig. 3). This implies that plants have no detectable amounts of 5hmC in DNA or very little, if any. This also suggests that 5hmC may not be used as a DNA demethylation substrate in plants by the DME/ROS1 family of proteins even though they have 5hmC excision activity.

4. Discussion

Here we report that the DME/ROS1 5mC DNA glycosylases excise 5mC oxidation derivatives such as 5hmC and 5caC. After 5mC or 5hmC excision, DME and ROS1 produce both 3'-PUA and 3'-phosphate via β - and δ -elimination processes, respectively (Fig. 1C). The 3'-end structures of these excision intermediates are converted to 3'-OH by AP endonuclease, allowing downstream DNA polymerase to incorporate unmethylated cytidine for the replacement of 5mC with cytosine (Fig. 1D). DME excises a significant amount of 5hmC although its excision rate is approximately 2.6-fold lower than that of 5mC (Fig. 2). The substrate competition assay also demonstrates that DME clearly has a preference for 5mC over 5hmC (Fig. 3). Despite our finding that DME/ROS1 proteins effectively remove 5hmC, we could not detect 5hmC from the plant DNA from TLC analysis (Fig. 4), implying that plants have no or very little 5hmC in the genome, if any, and may not utilize 5hmC as a primary 5mC-oxidation intermediate for DNA demethylation as in mammals.

Therefore, the 5hmC excision activity of DME/ROS1 may simply reflect a broad range of substrate specificity of this family of proteins, which is a common property of most DNA glycosylases [23,25,26]. The substrate specificity is determined by the recognition pocket of DNA glycosylase. The size of the electron cloud at C5 of 5hmC is larger than that of 5mC [27]. Therefore, 5hmC would need a larger recognition pocket than 5mC, which may explain why DME and ROS1 prefer 5mC to 5hmC. It was previously reported that ROS1 can excise several bases including 5-hydroxyuracil [24]. However, the finding that 5-hydroxymethyluracil was hardly excised by ROS1 [24] suggests the importance of the amine group at C6 of 5mC, possibly for thermodynamic stability of the target base via hydrogen bonding with amino acids inside the recognition pocket during base excision.

A recent study reported that a low but detectable amount of 5hmC was present in the Arabidopsis genome, which was demonstrated by immunoblotting methods using anti-5hmC antibody [28]. This raises the possibility of 5hmC serving as a DNA demethylation intermediate in plants, which can be processed by 5hmC-specific DNA glycosylases including DME/ROS1 family proteins. However, the approach that Yao et al. [28] used might not be sensitive enough to reflect the real base composition of plant genome due to non-specificity of antibody-based methods. In addition, no functional counterparts of TET-family enzymes or hydroxylating enzymes acting on 5mC have yet been identified in plants. Alternatively, 5hmC can be spontaneously produced by oxidative damages resulting from reactive oxygen species such as superoxide anions ($\bullet\text{O}_2^-$), hydroxyl radicals ($\bullet\text{OH}$), and hydrogen peroxide (H_2O_2) [29]. Therefore, a trace amount of 5hmC can be present in the plant genome with no aid of corresponding enzyme activities. It is still an intriguing question to be answered how plants and animals have evolved distinct DNA demethylation systems, in particular, the evolution of 5mC-specific DNA glycosylases in plants and the use of 5mC-oxidation derivative(s) intermediates mediating DNA demethylation in mammals.

To the best of our knowledge, DME/ROS1 proteins are the first reported DNA glycosylases that recognize and excise 5hmC from DNA. Even though its biological relevance in plants is still under question due to a lack of clear evidence for the existence of 5hmC in the plant genome, DME/ROS1 proteins have a promising potential for epigenome editing [30]. In particular, active DNA demethylation at specific targets combined with a transcription activator-like effector (TALE) or CRISPR system may allow transcriptional activation of the genes that are silenced by DNA methylation. A recent report demonstrated that TALE-TET1 fusion proteins successfully induced targeted DNA demethylation and gene expression [31]. Considering the fact that DME/ROS1 proteins remove both 5mC and 5hmC, the approach using these proteins can be used for epigenome editing in plants and mammals because 5mC is a universal silencing mark utilized in both systems.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by a grant from the Next-Generation BioGreen 21 Program (No. PJ009092), Rural Development Administration, and the Global Ph.D. Fellowship (NRF-2013H1A2A1034385, H.J.), National Research Foundation, Republic of Korea.

References

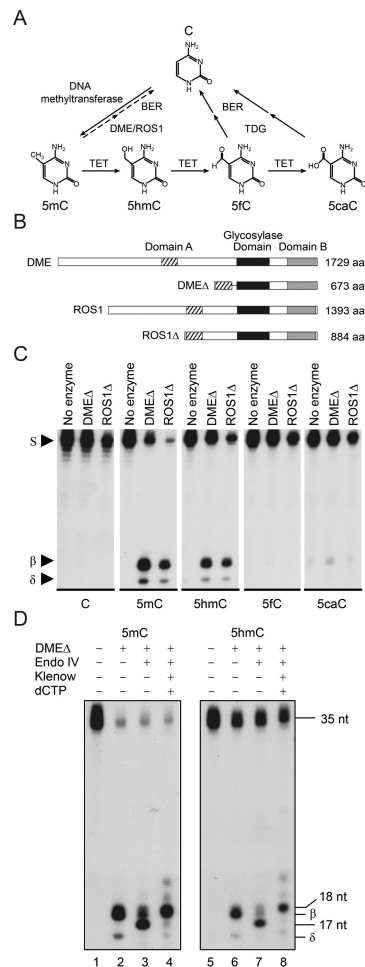
1. Bonasio R, Tu S, Reinberg D. Molecular signals of epigenetic states. *Science*. 2010; 330:612–616. [PubMed: 21030644]
2. Chan SW, Henderson IR, Jacobsen SE. Gardening the genome: DNA methylation in Arabidopsis thaliana. *Nat Rev Genet*. 2005; 6:351–360. [PubMed: 15861207]
3. Huh JH, Bauer MJ, Hsieh TF, Fischer RL. Cellular programming of plant gene imprinting. *Cell*. 2008; 132:735–744. [PubMed: 18329361]
4. Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet*. 2010; 11:204–220. [PubMed: 20142834]

5. Wu SC, Zhang Y. Active DNA demethylation: many roads lead to Rome. *Nat Rev Mol Cell Biol.* 2010; 11:607–620. [PubMed: 20683471]
6. Guo JU, Su Y, Zhong C, Ming GL, Song H. Hydroxylation of 5-methylcytosine by TET1 promotes active DNA demethylation in the adult brain. *Cell.* 2011; 145:423–434. [PubMed: 21496894]
7. He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, Sun Y, Li X, Dai Q, Song CX, Zhang K, He C, Xu GL. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science.* 2011; 333:1303–1307. [PubMed: 21817016]
8. Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y. Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature.* 2010; 466:1129–1133. [PubMed: 20639862]
9. Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science.* 2011; 333:1300–1303. [PubMed: 21778364]
10. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science.* 2009; 324:930–935. [PubMed: 19372391]
11. Xu Y, Wu F, Tan L, Kong L, Xiong L, Deng J, Barbera AJ, Zheng L, Zhang H, Huang S, Min J, Nicholson T, Chen T, Xu G, Shi Y, Zhang K, Shi YG. Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol Cell.* 2011; 42:451–464. [PubMed: 21514197]
12. Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PA, Rappsilber J, Helin K. TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature.* 2011; 473:343–348. [PubMed: 21490601]
13. Maiti A, Drohat AC. Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites. *J Biol Chem.* 2011; 286:35334–35338. [PubMed: 21862836]
14. Gehring M, Reik W, Henikoff S. DNA demethylation by DNA repair. *Trends Genet.* 2009; 25:82–90. [PubMed: 19144439]
15. Agius F, Kapoor A, Zhu JK. Role of the Arabidopsis DNA glycosylase/lyase ROS1 in active DNA demethylation. *Proc Natl Acad Sci.* 2006; 103:11796–11801. [PubMed: 16864782]
16. Gehring M, Huh JH, Hsieh TF, Penterman J, Choi Y, Harada JJ, Goldberg RB, Fischer RL. DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation. *Cell.* 2006; 124:495–506. [PubMed: 16469697]
17. Morales-Ruiz T, Ortega-Galisteo AP, Ponferrada-Marin MI, Martinez-Macias MI, Ariza RR, Roldan-Arjona T. DEMETER and REPRESSOR OF SILENCING 1 encode 5-methylcytosine DNA glycosylases. *Proc Natl Acad Sci.* 2006; 103:6853–6858. [PubMed: 16624880]
18. Ortega-Galisteo AP, Morales-Ruiz T, Ariza RR, Roldan-Arjona T. Arabidopsis DEMETER-LIKE proteins DML2 and DML3 are required for appropriate distribution of DNA methylation marks. *Plant Mol Biol.* 2008; 67:671–681. [PubMed: 18493721]
19. Penterman J, Zilberman D, Huh JH, Ballinger T, Henikoff S, Fischer RL. DNA demethylation in the Arabidopsis genome. *Proc Natl Acad Sci.* 2007; 104:6752–6757. [PubMed: 17409185]
20. Choi Y, Gehring M, Johnson L, Hannon M, Harada JJ, Goldberg RB, Jacobsen SE, Fischer RL. DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in Arabidopsis. *Cell.* 2002; 110:33–42. [PubMed: 12150995]
21. Mok YG, Uzawa R, Lee J, Weiner GM, Eichman BF, Fischer RL, Huh JH. Domain structure of the DEMETER 5-methylcytosine DNA glycosylase. *Proc Natl Acad Sci.* 2010; 107:19225–19230. [PubMed: 20974931]
22. Bhutani N, Burns DM, Blau HM. DNA demethylation dynamics. *Cell.* 2011; 146:866–872. [PubMed: 21925312]
23. Scharer OD, Jiricny J. Recent progress in the biology, chemistry and structural biology of DNA glycosylases. *Bioessays.* 2001; 23:270–281. [PubMed: 11223884]
24. Ponferrada-Marin MI, Roldan-Arjona T, Ariza RR. ROS1 5-methylcytosine DNA glycosylase is a slow-turnover catalyst that initiates DNA demethylation in a distributive fashion. *Nucleic Acids Res.* 2009; 37:4264–4274. [PubMed: 19443451]

25. Bulychev NV, Varaprasad CV, Dorman G, Miller JH, Eisenberg M, Grollman AP, Johnson F. Substrate specificity of Escherichia coli MutY protein. *Biochemistry*. 1996; 35:13147–13156. [PubMed: 8855952]
26. Dizdaroglu M, Laval J, Boiteux S. Substrate specificity of the Escherichia coli endonuclease III: excision of thymine- and cytosine-derived lesions in DNA produced by radiation-generated free radicals. *Biochemistry*. 1993; 32:12105–12111. [PubMed: 8218289]
27. Rangam G, Schmitz KM, Cobb AJ, Petersen-Mahrt SK. AID enzymatic activity is inversely proportional to the size of cytosine C5 orbital cloud. *PLoS One*. 2012; 7:e43279. [PubMed: 22916236]
28. Yao Q, Song CX, He C, Kumaran D, Dunn JJ. Heterologous expression and purification of Arabidopsis thaliana VIM1 protein: in vitro evidence for its inability to recognize hydroxymethylcytosine, a rare base in Arabidopsis DNA. *Protein Expr Purif*. 2012; 83:104–111. [PubMed: 22459921]
29. Loenarz C, Schofield CJ. Oxygenase catalyzed 5-methylcytosine hydroxylation. *Chem Biol*. 2009; 16:580–583. [PubMed: 19549596]
30. Voigt P, Reinberg D. Epigenome editing. *Nat Biotechnol*. 2013; 31:1097–1099. [PubMed: 24316647]
31. Maeder ML, Angstman JF, Richardson ME, Linder SJ, Cascio VM, Tsai SQ, Ho QH, Sander JD, Reyon D, Bernstein BE, Costello JF, Wilkinson MF, Joung JK. Targeted DNA demethylation and activation of endogenous genes using programmable TALE-TET1 fusion proteins. *Nat Biotechnol*. 2013; 31:1137–1142. [PubMed: 24108092]

Abbreviations

DME	DEMETER
5hmC	5-hydroxymethylcytosine
5mC	5-methylcytosine
ROS1	REPRESSOR OF SILENCING 1
TLC	thin layer chromatography

**Fig. 1.**

DME and ROS1 excise 5mC and 5hmC *in vitro*. (A) DNA demethylation pathways in mammals and plants. Plant-specific DNA demethylation steps are indicated with dashed arrows. (B) Structures of DME and ROS1 proteins along with their truncated versions used in the study. Three conserved domains – domain A (hatched box), glycosylase domain (solid box), and domain B (shaded box) – are necessary for the glycosylase activity [21]. (C) Base excision activity of MBP-DME and -ROS1 proteins for 5mC oxidation derivatives. Radiolabeled oligonucleotide substrate containing cytosine, 5mC, 5hmC, 5fC or 5caC was incubated with purified MBP-DME or MBP-ROS1 at 37 °C for 1 h. Reaction products were heat-denatured and separated on a 15% denaturing polyacrylamide gel. Oligonucleotide substrate (S) and β- and δ-elimination products (β, δ) are indicated to the left. (D) Reconstitution of DNA demethylation process after 5mC or 5hmC excision. Oligonucleotide substrate with 5mC (left panel) or 5hmC (right panel) was co-incubated with purified MBP-DME and Endonuclease IV (lanes 3 and 7). The reaction products were subjected to nucleotide extension with Klenow fragment (3' to 5' exo-) in the presence of dCTP (lanes 4 and 8). Positions of substrate (35 nt) and reaction products (17 nt, 18 nt) relative to β- and δ-elimination products (β, δ) were indicated to the right.

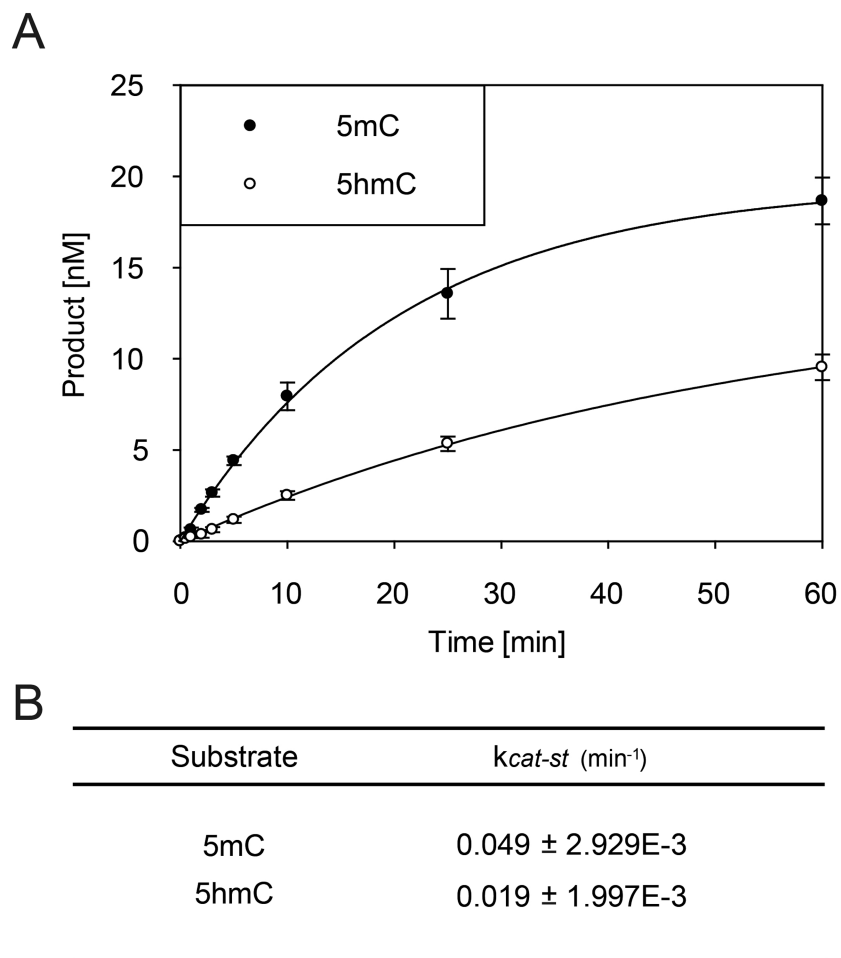


Fig. 2. Enzyme kinetics of DME for 5mC and 5hmC excision. (A) Oligonucleotide substrates (25 nM) containing 5mC or 5hmC were subjected to time-course reaction with MBP-DME (100 nM) and the reactions were terminated at 0.5, 1, 2, 3, 5, 10, 25, 60 min, respectively. Error bars represent standard deviations from three independent experiments. (B) The catalytic rate constants under single-turnover condition (k_{cat-st}) of DME for 5mC and 5hmC determined from Fig. 2A.

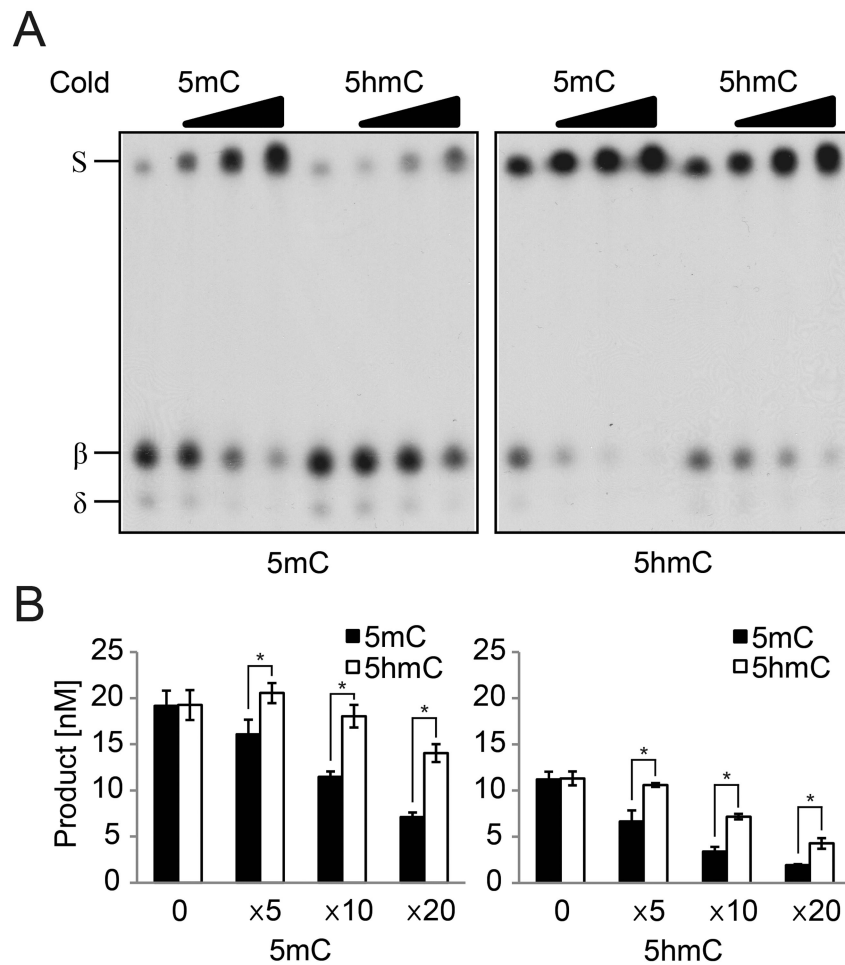


Fig. 3. Substrate competition assay of DME between 5mC and 5hmC. (A) Radiolabeled oligonucleotide substrates containing 5mC (left panel) or 5hmC (right panel) were incubated with MBP-DME in the presence of increasing amounts of 5hmC or 5mC cold competitors. The amount of competitors varied from 0 to 20-fold over radiolabeled substrate. Cold competitors are indicated at the top of each panel and radiolabeled substrates at the bottom. (B) Quantitation of the reaction products from the competition experiments (Fig. 3A). Error bars represent standard deviations from three independent experiments. A significant difference calculated from Student's t-test ($p < 0.05$) was denoted with an asterisk (*).

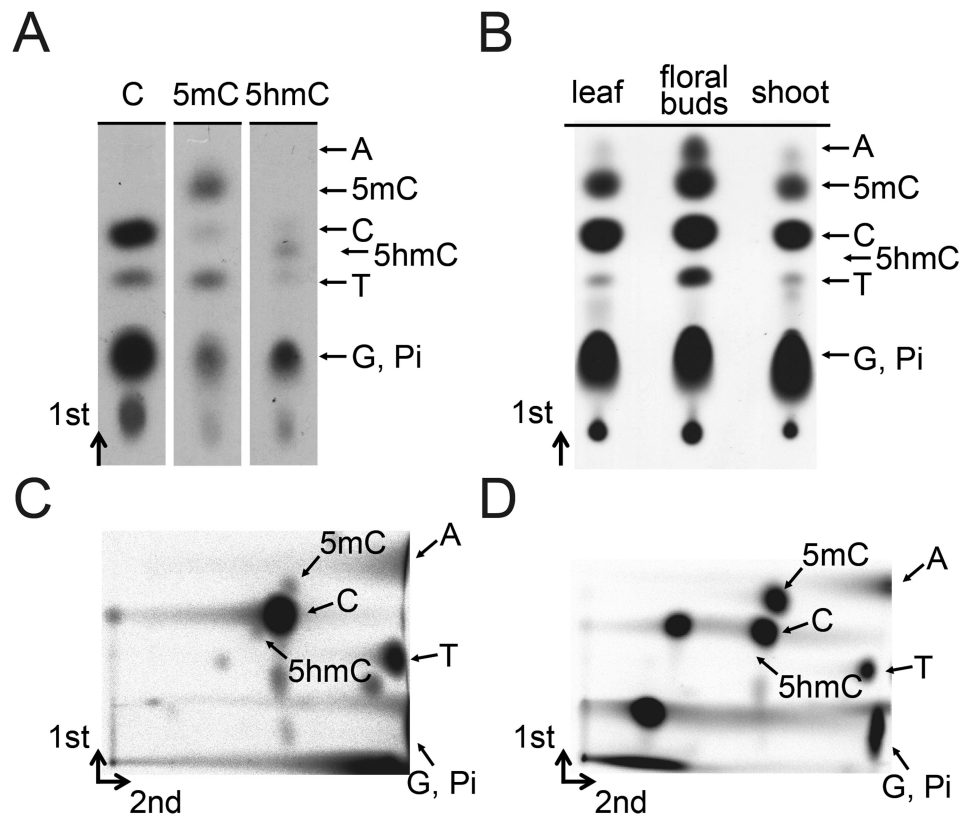


Fig. 4. *Arabidopsis* contains no detectable amount of 5hmC in the genome. (A) Synthetic DNA containing cytosine, 5mC, or 5hmC was digested with Taq^αI and radiolabeled with [γ -³²P]ATP as a control. The fragments were further digested to dNMPs and separated on a TLC plate. Note that 5-hydroxymethyl-dCMP was detected between the spots of dCMP and dTMP. (B) Genomic DNA isolated from leaves, floral buds, or whole plants was subjected to the same procedure as in Fig. 4A and separated on a TLC plate. (C, D) 2D-TLC analysis of samples derived from synthetic DNA with 0.2% each of 5mC and 5hmC (C) or from *Arabidopsis* floral buds (D). Spots corresponding to identifiable nucleotides were indicated with arrows according to the previous report [9]. See Suppl. Fig. 3 for additional 2D-TLC images.